ITS assignment 6

Noah Jensen (xdr622) & Julian Pedersen (rsk975)

October 2022

Forensics

Introduction

explain the problem

Attack and forensics

An invalid user alice in trying to ssh into our network using brute force, trying each port. later same IP tries users bob and eve on different ports too. Failed password for each attempt. time in log file shows that each ssh attempt happens within a second of the last, this implies automazation. the attacker changes user but continues the attack on trying each port. user tries several users, alice, bob, tom, eve, root. users that actually exist root and tom They are doing a brute force attack on random users with a password. We think the attackers have a password list and tries to find an open port and a user which owns the password. The attacker successfully enters tom's account with a right password on an open port. The login and immediately logout. The attacks stops for about 7 seconds, before starting again with a new invalid user.

We think the attackers obtained a list of password, and they are trying to match these passwords to users.

The attacker tries about 500 ports per user.

Response

STOP FIND THE BREACH AND REMOVE IT SWAP PASSWORDS Security analysis

Immediately ban the IP trying to connect. because the hackers got into Tom's account, we would change all the passwords of all users.

Future prevention

long term avoidance: ban IP's that fail to connect mutiple times within seconds. It is not normal behaviour. Maybe with a multiplicative-increasing time-out per unsuccessful connection attempt from a specific IP address. Also, simply adding extra time per connection attempt in all cases. implement a specification based IDS, now that we know how the attack works. maxtries automatic block of sus ips two factor auth on your ssh server.

```
grep -o -e "Failed password for.*from .............  port" auth.log
| sort | uniq -c
```

Short answers

Question 1

Compare forensics to incident response. What are the key similarities and differences?
Incident response is an immediate action against an ongoing attack to stop the attack, mitigate damages, reestablish security. Specific details of the attack and its effects are less relevant outside of those necessary to detect and thwart the attack itself.

Forensics is an after-the-fact action which involves collection, identification, examination, analysis, and reporting of data while preserving the integrity of the information and maintaining a strict chain of custody of data.[1] The custody chain is necessary to maintain its legitimacy for use in (among other things) the court-of-law.

Forensics is also used to cover other crimes than just attacks on a company or individual. For example, this can relate to an institution that might want to do forensics of a hard-disk to determine if the user had downloaded child pornography to use that information in the court-of-law.

Question 2

When a file is deleted, is it typically erased or not erased from the media?
Not erased, typically it is only the pointer to the file that is deleted, not the file itself. The data that is there is no longer ”protected” by a file system, allocation header/footers, or other types of data structuring, so it is essentially ”garbage” data and can be overwritten freely.

Question 3

In forensics, what is slack space?
When a file is deleted, it is not erased as explained above, but rather ready for reallocation. You normally allocate in discrete chunks, so if these chunks are not entirely used, leftover space is referred to as slack space, and can contain old garbage data from a previous allocation (for example, data left over from previous ”deletes” (see question 2)).

Question 4

In data privacy, what is the difference between data generalization and data randomization?

[1] https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-86.pdf, ES-1

Data randomization refers to changing attributes within a dataset to add uncertainty, while maintaining overall general distributions within the data set. For example, this can involve randomly changing all the ages in a dataset to be in a range of $[-3, +3]$ of the actual age (this specific example is "noise generation").

Data generalization refers to changing the *criteria* of a dataset (in scale and/or magnitude) so that they are common to a wider set of people. For example, this could be using age brackets $(0-9, 10-19$, etc.) instead of specific age values in the aforementioned dataset.